Research Methods and Data Analysis

Significance, Probability and Type I and Type II Errors

Significance

A null hypothesis predicts that there will be no effect, no correlation, no association; in short that nothing out of the ordinary will happen, or if it does then it's due to chance. The research hypothesis on the other hand predicts that there will be an effect, correlation or association and that it is not due to chance.

A finding is said to be ***significant*** if we can be fairly sure that it is unlikely to have occurred by chance. Significance therefore concerns the ***frequency*** of something occurring. For instance, if we were looking to see if memory recall of a list of words was improved by using a mnemonic, the null hypothesis would predict no improvement whilst the experimental hypothesis would predict an improvement. ***Any*** improvement that was unlikely to have occurred due to chance would be significant. Notice that significance does not depend upon how many more words were recalled but upon the occurrence of some improvement, no matter how small.

The main difficulty here is deciding how likely it is that a particular occurrence is due to chance or due to the variable we are investigating. This is decided on the basis of probability.

Probability

Probability is express by the letter ***p*** and as a decimal. Probability ***p*** is the likelihood of something occurring due to chance. The likelihood of picking a red card from a pack of playing cards is 1 in 2 or 1/2 and is written ***p*** = 0.50. If the likelihood of something is less than 1/2, ***p***<0.50. If the likelihood of something occurring is 1 in 4 or 1/4, like two tossed coins both landing heads, this is written ***p*** = 0.25. Similarly if the likelihood of something occurring is 1 in 20 or 1/20, like your name being drawn out of a hat of twenty names, this is written ***p*** = 0.05. It's fairly unlikely that your name will be drawn out of the hat.

As 1/20 is equivalent to 5/100, another way of saying this is that there's only 5% chance of your name being drawn. Therefore, you can be 95% confident your name will not be drawn.

If a coin were tossed 100 times you would expect 50 heads, 50 tails but you would not be too surprised if you got 51 heads, 49 tails. In fact, you would probably be in error if you claimed that something significant was happening. This would be making the mistake of falsely accepting the research hypothesis, but more of that later. You might not be too bothered if you got 60 heads, 70 heads, 75 heads ... but at what point would you be fairly sure something dodgy was going on? The convention in psychology is that we want to be 95% certain before we say that we're sure something interesting

other than ordinary chance is occurring. So the convention is that the probability of a particular event occurring due to chance (***p***) has to be as low as 5% or ***p*** = 0.05.

Of course, it is *possible* but very unlikely that a fair coin would come down heads 100 times out of 100 trials *by chance.* It is likely that one group of people will remember more than another group of very similar people but if you tested them 100 times you would be more than surprised if that group always scored better. You'd suspect that something significant had happened, for example that they were using a mnemonic. However, it is *possible* but very unlikely that this could occur due to chance, just like the coin landing heads up 100 out of 100 trials. If you claim that there was something significant going on you *could* be making an error.

Type I and Type II Errors

If you claim that something significant is happening when it is in fact due to chance, you are making a **Type I error**. This is when you make the mistake of *falsely rejecting the null hypothesis* and so *falsely accepting the research hypothesis*. This has occurred because you have allowed a margin of chance, in this case 5%. If you want to reduce the possibility of making such an error you could be stricter and adopt a more stringent level of ***p***, say ***p*** =0.01 or a 1% probability of the results occurring due to chance. The greater the margin of chance, the greater the risk of making a type I error. If you are unclear at this point, go back to the coin being tossed 100 times and getting 51 heads, 49 tails.

So why not use ***p*** = 0.01 or even ***p*** = 0.00001 to be sure of never falsely claiming that something was significant, never falsely rejecting the null hypothesis? Adopting too stringent a probability level is like saying that you will not be convinced that the coin is biased until it has been tossed and come down heads 100 times in 100 trials (***p*** = 0.01) or 100000 times in 100000 trials! (***p*** = 0.00001). If so, you would probably be missing the point that the coin is biased, claiming that nothing significant was happening until some poor fool had tossed the coin 100000 times and got 100000 heads! You might agree that this is making a mistake too. If you *falsely accept the null hypothesis* and *falsely reject the research hypothesis* you are making a **Type II error**.

If we set too stringent a significance level, such as 1%, we run the risk of making a type II error whilst reducing the risk of a type I error. Conversely, if we set too lax a significance level, such as 10%, we may well reduce the risk of making a type II error but increase the risk of making a type I error. So, 5%/***p*** = 0.05 seems just about a fine and dandy compromise to me!

Knowledge check

1. What does the null hypothesis predict about events?
2. If I win the lottery, is this due to chance or to something mysterious happening?
3. What does significance depend upon?

4. If I win the lottery, is it significant? If so, why? If not, why not?
5. What does ***p*** = 0.50 mean?
6. If you are 95% confident that your name will ***not*** be drawn out of a hat with 20 names in it, are you 95% confident in accepting the null hypothesis or 95% confident in accepting the research hypothesis?
7. If you are 5% confident it ***will*** happen, are you 5% confident in accepting the null hypothesis or in accepting the research hypothesis? Are you 5% confident in rejecting the null or in rejecting the research hypothesis?
8. If you had a pack of 52 playing cards, is it ***possible*** that the first 26 cards are red ***by chance***?
9. Define a Type I error and a Type II error, *in your own words*.
10. If the first 26 cards were all red, you'd probably want to say something significant was happening (micro-pk? clairvoyance?), but you ***could*** be making the error of claiming significance when something has happened by chance. What type of error is this?
11. If you were to insist that 25 cards had to be turned over before you would be confident in accepting the prediction that all the red cards are on top (the research hypothesis), are you being too stringent or too lax?
12. In the above example would you be making a type I or a type II error?
13. If the roof falls in when you read the number of this question and you claim that it's significant, do you run a greater risk of making a type I or a type II error?
14. If you wanted to avoid making such an error in this case, would you:
    a) accept the null hypothesis
    b) reject the research hypothesis
    c) leave the building at great speed
    d) all of the above?